



# In Search for Infall Gas in Molecular Clouds: A Catalogue of CO Blue-Profiles

Zhibo Jiang (江治波)<sup>1,2,3</sup> , Shaobo Zhang (张少博)<sup>1</sup> , Zhiwei Chen (陈志维)<sup>1</sup> , Yang Yang (杨旻)<sup>1,2</sup> , Shuling Yu (于书岭)<sup>1,2</sup> , Haoran Feng (冯浩然)<sup>1,2</sup> , and Ji Yang (杨戟)<sup>1</sup>   
the MWISP Group

<sup>1</sup> Purple Mountain Observatory, Chinese Academy of Sciences, Nanjing 210023, China; [zbjiang@pmo.ac.cn](mailto:zbjiang@pmo.ac.cn)

<sup>2</sup> University of Science and Technology of China, Hefei 230026, China

<sup>3</sup> Center for Astronomy and Space Sciences, Three Gorges University, Yichang 443002, China

Received 2022 August 18; revised 2023 February 1; accepted 2023 February 27; published 2023 June 9

## Abstract

We have started a systematic survey of molecular clumps with infall motions to study the very early phase of star formation. Our first step is to utilize the data products by MWISP to make an unbiased survey for blue asymmetric line profiles of CO isotopical molecules. Within a total area of  $\sim 2400$  square degrees nearby the Galactic plane, we have found 3533 candidates showing blue-profiles, in which 3329 are selected from the  $^{12}\text{CO}$  &  $^{13}\text{CO}$  pair and 204 are from the  $^{13}\text{CO}$  &  $\text{C}^{18}\text{O}$  pair. Exploration of the parametric spaces suggests our samples are in the cold phase with relatively high column densities ready for star formation. Analysis of the spatial distribution of our samples suggests that they exist virtually in all major components of the galaxy. The vertical distribution suggest that the sources are located mainly in the thick disk of  $\sim 85$  pc, but still a small part are located far beyond Galactic midplane. Our follow-up observation indicates that these candidates are a good sample to start a search for infall motions, and to study the condition of very early phase of star formation.

Supplementary material for this article is available [online](#)

*Key words:* stars: formation – ISM: clouds – catalogs – surveys

## 1. Introduction

Gravitational collapse of dense molecular cloud cores is a key step in the formation of stars (e.g., Shu et al. 1987). Presently due to the limitation of the observational facilities, it is difficult to obtain a motion picture showing a core in the process of gravitational collapse. However, line profiles can provide signatures of inward collapse motion. As theoretical modelings show (e.g., Myers et al. 1996), after a collapse happens, the center of the core becomes warmer due to the accumulation of gravitational energy while the outside envelope remains relatively cool. In the configuration of an inflowing core surrounded with a static envelope, the radiation transfer along the line of sight can be simplified by a two layer model for both the gas flowing away and toward the observer (see Figure 5 in Evans 1999). For the gas flowing toward the observer, the inner layer with higher excitation is nearer to the observer than the outer layer with lower excitation; for the flow-away gas, emission from the inner layer with higher excitation is absorbed by the outer layer with lower excitation lying closer to the observer. A double-peaked line profile with an absorption dip is produced by the above radiation transfer. The blue peak (gas flowing toward the observer) is stronger than the red peak (gas flowing away from the observer). This

effect is especially conspicuous for molecular transition lines with a suitable optical depth and critical density. This kind of profiles are commonly referred to as “blue-profiles.” They have been discussed by many researchers, and are regarded as an indicator of gas inflow motion (e.g., Zhou et al. 1993; Mardones et al. 1994; Myers et al. 1996; Mardones et al. 1997), or infall signature. To distinguish a blue-profile from multi-component emissions in the line of sight that may also show double-peaks, one needs an optically thin line, whose profile is single-peaked without a self-absorption feature. The peak position of the optically thin line should be close to the self-absorption dip between the blue and red peaks.

Observational studies of the blue-profiles started in early 1990s (Zhou et al. 1993; Mardones et al. 1994) and are seen accelerated recently. Up to date, the studies can be classified into two categories. One is case study, i.e., detailed observations toward some particular star-forming complexes to study the physical conditions of regions showing infall signature (e.g., Zhou et al. 1993; Zhu et al. 2011; Ren et al. 2012; Mayen-Gijon et al. 2014; Yuan et al. 2018). The other is the deliberated search, i.e., systematical searches for infall signatures in deliberately selected samples of different properties, such as low-mass protostellar objects (Mardones et al. 1997), high-mass star-

forming regions (Wu & Evans 2003; Klaassen & Wilson 2007; Wu et al. 2007; Sun & Gao 2009; Jin et al. 2016), high-mass protostellar objects (Yoo et al. 2018), high infrared extinction clouds (Rygl et al. 2013), massive star-forming cores (Sun & Gao 2009; He et al. 2015), early dense cores (Calahan et al. 2018), and extended green objects (Chen et al. 2010), etc.

Above studies have presented us with diverse hints on the initial phase of star formation, e.g., when the gravitational collapse starts and ends, whether high- and low-mass stars start their life in similar ways, and the timescale of mass-infall phase. However, a comprehensive study on the infall phase is necessary to understand fully the process how dense molecular cores turn into stars, and to contribute to solve the long-standing problem of the origin of the stellar initial mass function. Previous studies mainly utilized known targets that have shown star-forming activities as a start to look into. This kind of researches will inevitably introduce bias if a statistical study is carried out upon them. We therefore do the work in an opposite direction, i.e., by merely looking for infall signatures in the galaxy without any other assumption, we study the physical properties and star-forming activities where the signatures happen. The ongoing Milky Way Imaging Scroll Painting Project (MWISP, Su et al. 2019), a large scale survey of CO ( $J = 1-0$ ) lines in the northern galaxy, provides an excellent opportunity to start up such a kind of work.

Our project is to set up a sample of CO blue-profiles using the MWISP data by blind search, and refine it by further observations using some more infall-sensitive lines (such as  $\text{HCO}^+$  lines). The aim of this project is to obtain a comprehensive sample of infall candidates with a high confidence level. This paper is to present the very first result of our project, a preliminary catalog of CO blue-profiles.

The structure of this paper is arranged in this way: in Section 2 we give a brief introduction to data preparation, strategy of blind search for blue-profiles; in Section 3 we present the main catalog and the associated line profiles; in Section 4 we discuss the distribution and parameter space of sources; and Section 5 summarizes our results.

## 2. Method

### 2.1. Data

All data used in this work are based on the MWISP data products. A detailed description of the data acquisition, quality control and archiving scheme can be found in Su et al. (2019). The archive data are in unit of “cells”, the main region of which is an area of  $30' \times 30'$ . According to the survey strategy, the outskirts of each cell are observed with the 9-beam Superconducting Spectroscopic Array Receiver (SSAR, Shan et al. 2012) less than the main region. Consequently, the rms noise levels are higher in the edge part than in the central. Therefore, a combination with the neighboring cells is necessary to achieve a uniform noise level. At this step, we intentionally extend the

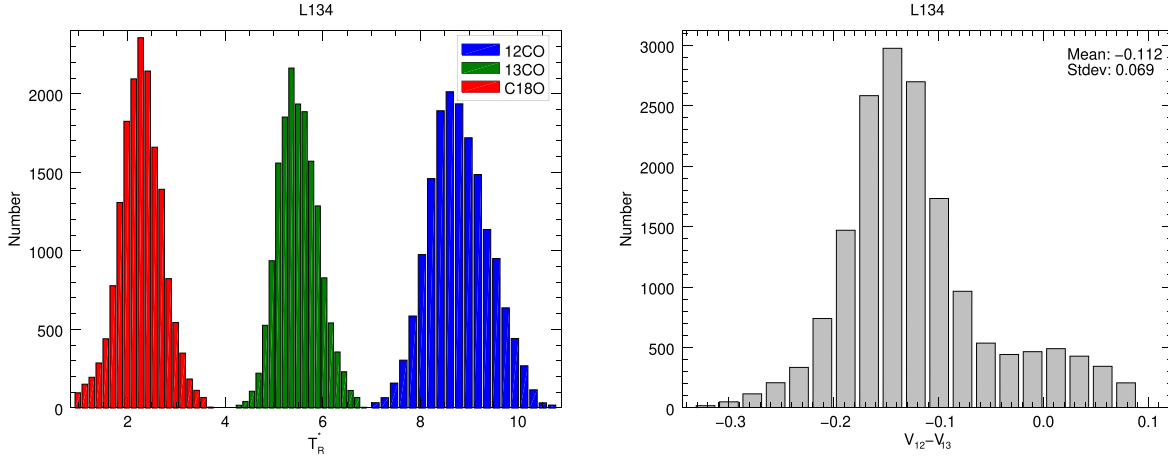
area of each cell to some extent (generally by 0.5 on each side) to avoid possible edge effects. We then cut off unnecessary velocity channels ( $|V_{\text{LSR}}| \geq 200 \text{ km s}^{-1}$ ) to minimize the unnecessary computation, and possible unreal detections due to untrue signals which might be a problem in the manual check step. The data reduced as such are ready for machine work.

Because our work relies on the comparison among different lines, it is an essential demand that the observation facility should be stable on a large timescale. Fortunately, MWISP has a scheme that immediately before and after the observation to each cell, observations toward a “standard source” were made to monitor the system performance. As an example, in Figure 1 we show the distributions of  $T_{\text{MB}}$  and  $V_{\text{LSR}}$  difference of L134, one of the “standard sources.” We choose this source as an illustration because its line-widths are relatively small to allow accurate estimation of the central velocity. The gross data were collected from 2011 November to 2017 December in all three lines. The  $1\sigma$  variations of  $T_{\text{MB}}$  are 0.66 K, 0.47 K, and 0.47 K, or 7.4%, 8.4%, and 20% on relative scale for the three lines, respectively. This suggests the system was reasonably stable within the period of seven years. The radial velocities with respect to local standard of rest (LSR) are also checked within the period. As shown in Figure 1,  $V_{\text{LSR}}^{12\text{CO}}$  is roughly less than  $V_{\text{LSR}}^{13\text{CO}}$  by  $\sim 0.15 \text{ km s}^{-1}$  (blueshifted). Martin & Barrett (1978) showed a slight velocity shift between  $^{12}\text{CO}$  and  $^{13}\text{CO}$ , which is consistent with this result. This difference could be intrinsic or arise in the observations. Considering that the  $V_{\text{LSR}}$  differences between  $^{12}\text{CO}$  blue peak and  $^{13}\text{CO}$  center in our selected sample are mostly significantly greater than that value, this discrepancy, even if could arise due to the instrument and observation, does not affect our result too much.

### 2.2. Search Strategy

As stated above, two lines, one being optically thick and the other optically thin, are necessary to discriminate between blue-profiles and multi-components. We use the  $^{12}\text{CO}$  with  $^{13}\text{CO}$  (hereafter Pair-1), and  $^{13}\text{CO}$  with  $\text{C}^{18}\text{O}$  (hereafter Pair-2) as two line pairs to do the work, i.e., we arbitrarily assume  $^{13}\text{CO}$  as an optically thin line and  $^{12}\text{CO}$  as the optically thick to search for blue-profiles, and then do it again with the other line pair. This scheme is reasonable because the  $^{13}\text{CO}$  line could be optically thin in some cases while thick elsewhere. We note that in the former case, the optically thin assumption of  $^{13}\text{CO}$  lines, which might be not true, does not affect the judgement of  $^{12}\text{CO}$  blue-profiles, but might miss some candidates if  $\text{C}^{18}\text{O}$  emissions are not detected.

Similar to the scheme of the MWISP survey, the search regions for machine work are split into cells. Then the automatic search is carried out pixel by pixel within each cell. As stated in Section 2.1, the area for each cell has a 0.5 overlap with the surrounding cells, redundant detections are inevitable, which are cleaned out at later steps.



**Figure 1.** (Left) The distribution of main-beam temperatures of L134 on a timescale of seven years (2011–2017). The blue, green and red histograms represent  $^{12}\text{CO}$ ,  $^{13}\text{CO}$  and  $\text{C}^{18}\text{O}$  peak intensities, respectively. (Right) The distribution of velocity difference between  $^{12}\text{CO}$  and  $^{13}\text{CO}$  in the same period. The figure shows the radial velocities of  $^{12}\text{CO}$  are overall less than that of  $^{13}\text{CO}$  by  $\sim 0.15 \text{ km s}^{-1}$  (about one channel). A secondary peak at  $\sim 0.1 \text{ km s}^{-1}$  is found, mainly caused in the the 2011 observation season. The results of all nine beams are included.

### 2.3. Automatic Search

Up to date, MWISP has been accomplished the region  $l = [12^\circ, 230^\circ]$ , and  $b = [-5^\circ.25, +5^\circ.25]$ , with some coverage at  $l < 12^\circ$ , covering a total area of  $\sim 2400$  square degrees, and produced  $\sim 10^8$  spectra. It is impossible to deal with such a large amount of data manually. Initial automatic searches have to be conducted to minimize the manual tasks. At this step for each pixel, we assume  $^{13}\text{CO}$  to be optically thin for Pair-1 and  $\text{C}^{18}\text{O}$  for Pair-2. First, we decomposed the optically thin lines into components with a one-dimensional watershed algorithm. Each component is Gaussian-like with a peak exceeding the nearby baseline or dip by over 5 rms of the whole spectrum. Then the first and second moments were extracted for each component as center velocity and velocity dispersion. They were aligned with the optically thick counterpart to cut out segments of profiles to be selected. Two methods were then adopted to find profiles with double peaks.

Method one: we select optically thick profiles fulfilling the following criteria.

1. A peak should appear on the blue side of central velocity. Such peak must exceed the nearby dip or baseline by 3 rms, and away from the center velocity by more than velocity resolution.
2. The peak on the blue side should be higher than the profile on the red side by at least 3 rms. If a peak appears on the red side, the component is labeled with “double-peak,” or a “shoulder” label will be assigned.

Method two: Gaussian fittings with two components were conducted to the optically thick profiles, and then results fulfilling the following criteria are selected.

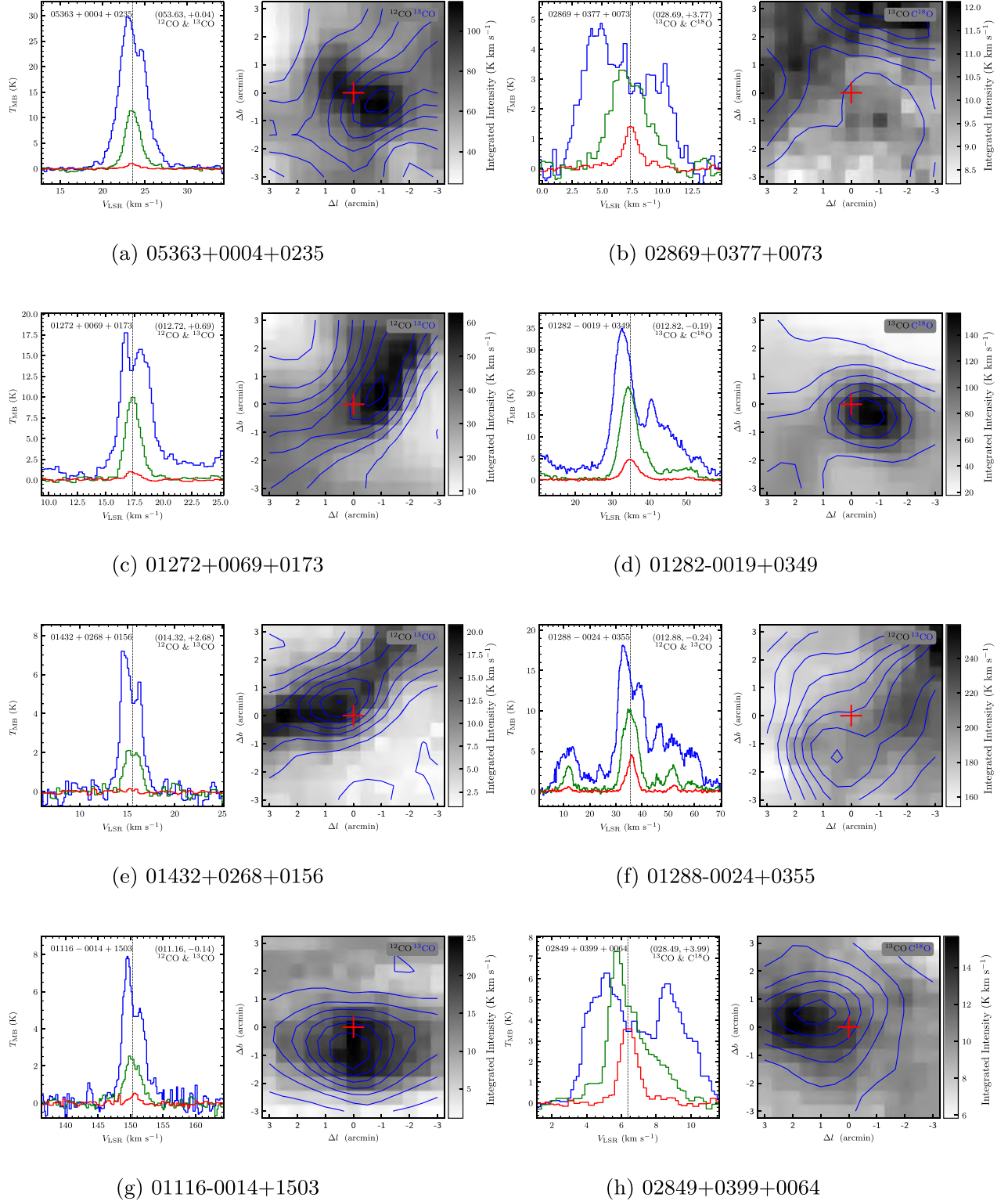
1. Two components should reside on both sides of the center velocity, and away from it by more than velocity resolution.
2. The blue side Gaussian component must have a higher integrated intensity and peak.

### 2.4. Manual Check

First look at the spectra of the candidates resulted from the automatic search suggests that many of them are not really blue-profiles, or hardly to say that. Therefore, further manual check is necessary to refine the candidate list. Picking up the candidates of likely blue-profiles is a hard work. We wrote some codes with graphic user interface to assist task. Then the pick-up work is only a couple of mouse clicks.

By checking the spectra of the candidates and comparing with the numeric simulations (see Section 4.1), the candidates are rejected according to the following general rules:

1. Spectra of either optically thick or thin lines are noisy: In the real data, the situation is much more complicated than the simulation due to the introduction of noises. Further more, some candidates appear at the edge of a certain cell where the noise level is higher than those of central areas. Though this situation has been considered in the automatic search, there are still cases that the machine would misjudge whether a peak is real or noise-induced, especially for optically thin lines.
2. Two peaks of the optically thick line are far away (e.g.,  $V_{\text{red}} - V_{\text{blue}} > 10 \text{ km s}^{-1}$ ): This is only empirical since we seldom find blue-profiles with very large peak separation in the literature. The line width (Full Width at Half



**Figure 2.** The spectra (left) and gas distributions (right) of the candidates. On the left panels, the red, green and blue lines represent  $\text{C}^{18}\text{O}$ ,  $^{13}\text{CO}$  and  $^{12}\text{CO}$  emissions, respectively. To enhance s/n ratios of the lines, the spectra are smoothed with median filter for each three velocity channels. For the same reason, the  $\text{C}^{18}\text{O}$  spectra are also smoothed with  $3 \times 3$  spatial pixels. The dashed line indicates the system velocity. On the upper-right corners we indicate the Galactic coordinates ( $l$ ,  $b$ ), and the line-pairs used in the searching procedure. On the right panels show the integrated intensity map of the optically thin lines (contours) overlaid on those of the optically thick lines (gray scale). The red pluses indicate the positions where the blue-profiles are detected.



Maximum, FWHM) of the optically thin line is too large, i.e.,  $>10 \text{ km s}^{-1}$ .

3. Multi-peaks or flattened top in the optically thin line: It is usually very difficult to judge whether a blue-profile of the optically thick line comes from a single or multiple components if there exist multi-peaks or flattened top in the assumed optically thin line. Though there are cases that the peaks of the optically thin lines do not coincident with those of the optically thick lines, we suppose they still could arise from multi-components with very complicated physical conditions. However, the candidates with one situation are retained: for Pair-1 selected candidates, the  $^{13}\text{CO}$  line also shows a blue-profile while the  $\text{C}^{18}\text{O}$  emission is not detected. The two peaks of the  $^{13}\text{CO}$  line are between those of  $^{12}\text{CO}$  line. In this case it is likely that  $^{13}\text{CO}$  line is also optically thick, but the absorption is fainter than that of the  $^{12}\text{CO}$  line. In such a case self-absorption still exists to show a blue-profile, but the peak separation is smaller than that of the  $^{12}\text{CO}$  line.

### 2.5. Cleaning Candidates

The manual check is a tiresome work since the output of the automatic search returns hundreds of thousands of candidates. In the manual check step we cautiously relax the above criteria in order to avoid loss of some reliable candidates. After the manual checks, we make the Gaussian fittings to the line profiles. Since the fittings need human interactions in deciding parameters such as fitting ranges, line profiles of the candidates are checked again and some candidates are rejected at this step.

After the manual check, candidates with similar positions and velocities are grouped together. The similarity of two candidates is defined as the following two conditions are satisfied:

$$\Delta p = \sqrt{\Delta l^2 \cos^2 b + \Delta b^2} \leq \Delta p_0; \Delta V_c \leq \Delta V_0$$

where  $\Delta l$ ,  $\Delta b$  and  $\Delta V_c$  are the differences of the Galactic coordinates and central velocities of two candidates;  $\Delta p_0 = 1.0$  and  $\Delta V_0 = 0.2 \text{ km s}^{-1}$  in this work being based on the spatial and spectral resolutions of the MWISP products. For each group the sources are checked by eyes and those showing blue-profile most significantly are selected.

Lastly, in a few cases blue-profiles are detected in both Pair-1 and Pair-2, we merge them to the Pair-2 group since the physical parameters (see Figure 2(e)), such as column densities and system velocities, could be more correctly estimated due to the fact that the optically thin lines are single-peaked.

## 3. The Catalogue

Finally, a total of 3533 sources are registered as possessing blue-profiled lines, out of which 3329 sources are selected from Pair-1, and 204 sources are selected from Pair-2. In Table A1

we present a subset of the candidates for illustration, while the whole catalog is presented in the appendix.<sup>4</sup> For each candidate we assign a name code, which is presented in Column 2. Name codes are based on the positions and velocities of the candidates, i.e., the  $(l, b)$  codes are reserved to the second decimal places and  $v$  to the first decimal places, and decimal points are removed from the final code. As an example, if  $(l, b, v) = (123.450000, 1.230000, 123.4500)$ , the code is assigned as “12345+0123+1235.” Columns 3 and 4 give the Galactic coordinates of sources. Column 5 gives the spatial association to the objects from some well known catalogs, such as IRAS PSC, NGC, the RAFGL catalog, the Sharpless catalog, and Lynds’ dark nebular catalog, etc. We also checked some other catalogs such as  $\text{H}_{\text{II}}$  candidates (Anderson et al. 2014; Yan et al. 2018), and the MSX dark cloud catalog (Simon et al. 2006). In Column 6 we indicate whether our candidates show blue-profiles with other tracers in the literature. The associations in Columns 5 and 6 are shown in codes to save the table space, and noted as table footnotes. In Column 7 we give the distances of the sources. The distance of a source is obtained using the online distance calculator<sup>5</sup> which is based on the trigonometric parallax measurements of Galactic maser sources (Reid et al. 2016, 2019) and assumes all molecular clouds are located on spiral arms. Using a Bayesian approach, this tool is believed to be a major improvement of the kinematic distance estimate of Galactic molecular clouds. For a few sources that are near the tangent direction ( $l \sim 90^\circ$ ), the calculator does not give proper distances.

The second half of the table presents the derived parameters of the candidates. Columns 8 & 9 ( $V_c, V_b$ ) are the characteristic velocities of sources.  $V_c$  represents the “central” velocities, which are derived from Gaussian fittings to the optically thin lines.  $V_b$  represents positions where blue peaks are found, which are largely obtained from double-component Gaussian fittings. However, in some cases when Gaussian fittings do not converge, we use the peak positions by eyes. Column 10 presents the skewness parameter,  $\delta V = (V_{\text{thick}} - V_{\text{thin}}) / \Delta V_{\text{thin}}$ , as suggested by Mardones et al. (1997).

Columns 11–13 ( $T_1, T_2, T_{\text{ex}}$ ) are related temperatures of the sources.  $T_2$  represents the peak main beam temperature of the optically thin line.  $T_1$  represents the expected main beam temperature of the optically thick line. Since the line profile shows self-absorbed feature, the “expected” peak is derived by Gaussian fitting to the whole profile while excluding the portion that shows self-absorption, i.e., we use the waist and foot portions in the fittings. Though the fitting portions have been carefully adjusted, we caution that in some cases the derived values of  $T_1$  are sensitive to the fitting portions adopted, so that the errors could be very large—they may reach

<sup>4</sup> The full table can also be found at <https://www.scidb.cn/en/detail?dataSetId=bc1c08c6e7ba426a9a07b8bad69b2ffb>.

<sup>5</sup> <http://bessel.vlbi-astrometry.org/node/378>

the values of  $T_1$  themselves in some extreme cases depending on the environments of individual sources, e.g., when multi-components coexist with the blue-profiles. However, in most cases, we expect the uncertainties being less than 20%.  $T_{\text{ex}}$  is the excitation temperature calculated from  $T_1$  (Mangum & Shirley 2015):

$$T_{\text{ex}} = T_0 \left[ \ln \left( 1 + \frac{1}{\frac{T_1}{T_0} + \frac{1}{\exp(T_0/T_{\text{bg}}) - 1}} \right) \right]^{-1}$$

where  $T_0 = h\nu_0/k = 5.53$  K for Pair-1, or 5.29 K for Pair-2,  $T_{\text{bg}}$  ( $= 2.73$  K) is the background temperature of the universe. Here we assume the filling factor being close to unity. We use  $^{13}\text{CO}$  to estimate the excitation temperatures of Pair-2 sources because in many cases  $^{12}\text{CO}$  profiles are very complicated so that we cannot obtain reliable “expected” main beam temperatures. In such cases  $T_{\text{ex}}$  might be underestimated since the  $^{13}\text{CO}$  lines are not optically thick enough.

The column densities of  $\text{H}_2$  (Column 14) are estimated from the integrated intensities of the optically thin lines assuming local thermal equilibrium (LTE). Under such an assumption, the column densities of  $^{13}\text{CO}$  or  $\text{C}^{18}\text{O}$  are calculated by the following formula (Mangum & Shirley 2016):

$$N(^{13}\text{CO}|\text{C}^{18}\text{O}) = f_1 \times f_0 \times \frac{(T_{\text{ex}} + 0.88)\exp(\frac{T_0}{T_{\text{ex}}})}{\exp(\frac{T_0}{T_{\text{ex}}}) - 1} \times \frac{\int T_{\text{MB}} dv \quad (\text{K km s}^{-1})}{f(J_\nu(T_{\text{ex}}) - J_\nu(T_{\text{bg}}))} \quad (\text{cm}^{-2})$$

where  $f_0$  is approximately  $2.48 \times 10^{14}$  for both  $^{13}\text{CO}$  and  $\text{C}^{18}\text{O}$ ,  $T_0$  equals 5.29 for  $^{13}\text{CO}$  and 5.27 for  $\text{C}^{18}\text{O}$ ,  $f$  is the beam filling factor which we assume to be 1.0 in this work, and  $J_\nu(T) = T_0/(\exp(T_0/T) - 1)$ . The optical depth correction factor  $f_1 = \tau/(1 - \exp(-\tau))$ . Where  $\tau$  is calculated by:

$$\tau = -\ln[1 - \frac{T_2}{f(J_\nu(T_{\text{ex}}) - J_\nu(T_{\text{bg}}))}]$$

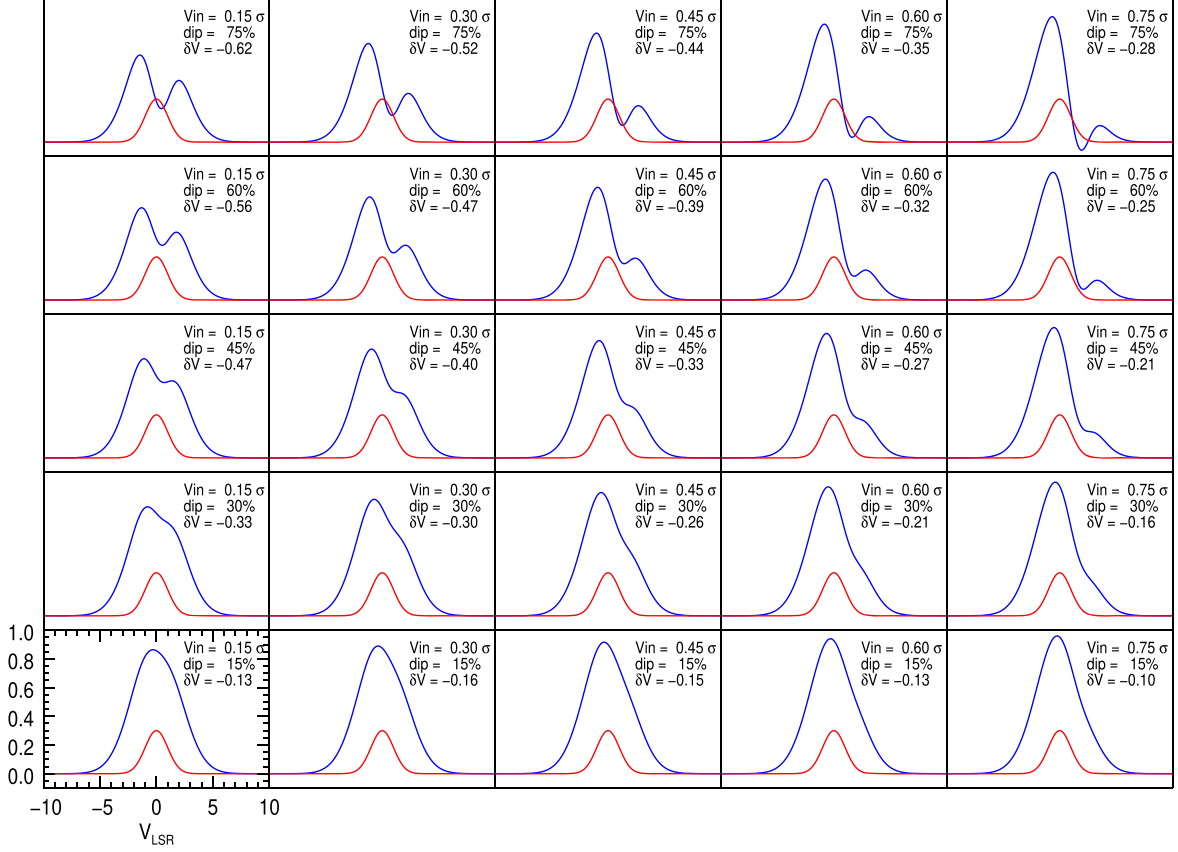
The total column densities of molecular hydrogen,  $\text{N}(\text{H}_2)$ , are obtained assuming the CO abundance ratios,  $\text{N}(\text{H}_2) \approx 3.9 \times 10^5 \text{N}(^{13}\text{CO})$  or  $\text{N}(\text{H}_2) \approx 3.0 \times 10^6 \text{N}(\text{C}^{18}\text{O})$  (Areal et al. 2019), which are derived from the infrared dark cloud and close to those from star-forming regions (Pineda et al. 2008). Several factors affect estimations of  $\text{N}(\text{H}_2)$ : (1) assumption of LTE which may not be true if the sources are actually in the process of infall; (2)  $T_{\text{ex}}$ s estimations; (3) abundances of  $^{13}\text{CO}$  and  $\text{C}^{18}\text{O}$  relative to those of molecular hydrogen change from region to region (Heyer & Dame 2015); (4) beam filling factors for the sources are not always close to unity. Therefore, it is difficult to estimate the overall

uncertainties of  $\text{N}(\text{H}_2)$ . In the last column (Column 15) we indicate the line pairs adopted in searching for the blue-profiles.

Figure 2 shows some examples of the spectra (left) and the relative location (right) where blue-profiles are detected. Again, for the limitation of the main body, the whole set of the figures is given in the appendix. Figure 2(a) (05363+0004+0235) shows an ideal blue-profile in  $^{12}\text{CO}$  emission with virtually symmetric profiles both in  $^{13}\text{CO}$  and  $\text{C}^{18}\text{O}$  lines, the blue-profile happens at the saddle point between two local maxima (right panel); Figure 2(b) (02869+0377+0073) shows a candidate selected from Pair-2. The  $^{12}\text{CO}$  line of this source shows complicated peaks; Figure 2(c) (01272+0069+0173) shows a blue-profile selected from Pair-1, but the red side of the  $^{12}\text{CO}$  line is probably contaminated by another component. The  $\text{C}^{18}\text{O}$  line, weak as it is, also suggests such a possibility. However, we could recognize this candidate since the contamination is not too serious to affect our judgement; Figure 2(d) (01282-0019+0349) shows a candidate that does not show clear blue-profile, but the  $\text{C}^{18}\text{O}$  and  $^{13}\text{CO}$  peaks are shifted redward with respect to the  $^{12}\text{CO}$  peak. Such a configuration happens when infall velocity is relatively large while the optical depth is moderate (see the lower-right panel of Figure 3). Figure 2(e) (01432+0268+0156) shows an example that both  $^{12}\text{CO}$  and  $^{13}\text{CO}$  lines exhibit blue-profiles, with the peak separation of  $^{12}\text{CO}$  greater than that of the  $^{13}\text{CO}$  line, while  $\text{C}^{18}\text{O}$  is marginally detectable. The blue-profile happens at virtually the central part of a dense clump. The examples presented here are typical cases in our candidates.

In Figure 2 we also present some extreme cases, which might be of interests to readers. Figure 2(f) (01288-0024+0355) shows a source with wide profiles, even for the optically thin lines ( $^{13}\text{CO}$  and  $\text{C}^{18}\text{O}$ ), up to FWHM  $\sim 6.6$  and  $4.0 \text{ km s}^{-1}$ , respectively. Figure 2(g) (01116-0014+1503) shows a target that has a very large system velocity,  $\sim 150 \text{ km s}^{-1}$ . The kinematic distance being estimated  $\sim 8.0$  kpc, it is one of the most distant object in the catalog. Interestingly, this source presents blue-profiles both in  $^{12}\text{CO}$  and  $^{13}\text{CO}$  lines. Figure 2(h) (02849+0399+0064) shows a source that is rarely seen in the Galactic molecular clouds, i.e., the  $^{13}\text{CO}$  line intensity is stronger than that of the  $^{12}\text{CO}$  line. Aside of strong self-absorption of  $^{12}\text{CO}$ , there could be also a possibility of unusual  $^{13}\text{CO}$  abundance and excitation environment.

In the right panel of each subset, we show the integrated intensity map of individual candidate. An interesting fact that should be noted here is that the infall-signatures do not always coincide with the clump centers (the points with the highest column densities). It would be interesting to ascertain whether this is a real fact or simply an observational effect, which might be helpful in our understandings of the initial phase of star formation.



**Figure 3.** One example of the numerical experiments with  $f_{\sigma} = 0.5$ . The parameters are shown at the upper-right corner of each frame. The blue lines represent the self-absorbed optically thick lines showing blue-profile lines, and the red lines indicate the un-absorbed optically thin lines for comparison. We also present  $\delta V$  suggested by Mardones et al. (1997) for different cases.

## 4. Discussion

### 4.1. Numerical Experiments

In order to see what a blue-profiled line appears in different environments, and to assist us to identify infall candidates from a large number of outputs from the machine work, we have performed a series of numerical experiments. The experiments tried to simulate cores with inflowing gas, but with very simple models, e.g., fixed infall velocities, optical depths and line widths. Temperature and density profile variations are not considered in the experiments. To do this, we generate an emission line of Gaussian profile centered at  $0.0 \text{ km s}^{-1}$ , and  $\sigma = 2.0 \text{ km s}^{-1}$ , which has no physical meaning other than a scale in the velocity dimension, and upon which many other parameters are assumed. The peak value is normalized to unity. A number of absorption lines, which are also of Gaussian profiles but with negative values, are put upon the emission line to mimic the absorption. The centers of the absorption lines are shifted a little to the red by fractions of  $\sigma$ , i.e.,  $V_{\text{in}} = f_{dV}\sigma$ . Here we use  $V_{\text{in}}$  to designate the central velocity shift of the absorption feature because it is somehow related to the infall

velocity in the reality. The negative maxima are assumed to be a few percent of unit ( $\text{dip} = f_{\text{dip}}$ ) because the optical depths of infalling gas are expected to be optically thinner than the background emission. In the reality, the widths of the absorption lines are expected to be narrower than those of the emission lines, and so that are set to some fractions of  $\sigma$  ( $\sigma_{\text{abs}} = f_{\sigma}\sigma$ ).

In Figure 3 we show some results of the experiments with  $f_{\sigma} = 0.5$  and various  $f_{dV}$ 's and  $f_{\text{dip}}$ 's. The parameters are shown at the upper-right corner of each frame. Though we also have done the experiments by varying  $f_{\sigma}$ , the results do not show very much difference and are not shown in this paper. From Figure 3 we see that the observability of a blue-profile line strongly depends on the strength of the absorption ( $f_{\text{dip}}$ ), but less sensitive to the velocity shift ( $V_{\text{in}}$ ). For  $f_{\text{dip}} \sim 0.15$ , approximately corresponding to the optical depth  $\tau \sim 0.15$ , a blue-profile is hardly seen even in our noise-free experiments and is thus not likely to be detected in real observations. In the cases of  $f_{\text{dip}} \sim 0.3$  ( $\tau \sim 0.36$ ), red shoulders are seen clearly. In such cases real detections are mostly missed because our auto-search algorithm searches only for double peaks. The blue-

profiles become observable when  $f_{\text{dip}} \sim 0.45$  ( $\tau \sim 0.6$ ) with a small velocity shift but are still unclear at large velocity shifts. For nearly all velocity shifts, the signatures are clearly observable at  $f_{\text{dip}} \sim 0.6$  ( $\tau \sim 0.9$ ).

At the corner of each frame, we also present the dimensionless skewness parameter,  $\delta V$ , suggested by Mardones et al. (1997) as an indicator of the significance of blue asymmetry. As can be seen from Figure 3,  $|\delta V|$  generally increases as  $f_{\text{dip}}$  but decreases as  $f_{\text{dip}}$ . In fact, our experiments suggest the absolute value of  $\delta V$  decreases also with  $f_{\sigma}$ .

#### 4.2. Parameter Distributions

Up to date, the MWISP project covers an area of  $\sim 2400$  square degrees toward the northern Galactic plane. Though the sample is still quite far from complete, due to the fact that the blue-profiles are recognized without any other presumption, the blind search strategy enables us to avoid bias toward known star-forming regions. This kind of work will provide a better approach to explore parametric space of the infalling molecular clumps. In the following we discuss the parametric distributions of our samples.

In Figure 4(a) we present the distributions of central velocities of the optically thin lines. In the main frame we show the pair-separated distributions while the overall distribution is shown in the inset, and the rest figures are treated in the same way. The distribution is generally in Gaussian style peaked around  $5 \text{ km s}^{-1}$ , and shows excess in both wings. A large amount of sources gather within  $|V_{\text{LSR}}| \leq 10 \text{ km s}^{-1}$ , suggesting that our sources are mainly located close to the Sun. We notice here, though not shown in the figure, the smallest and largest central velocity are  $-75.5$  and  $150.3 \text{ km s}^{-1}$ , respectively. An interesting feature seen from Figure 4(a) is that Pair-1 selected sources have lower radial velocities (with a median value of  $\sim 6.8 \text{ km s}^{-1}$ ) than Pair-2 selected sources (median  $\sim 7.9 \text{ km s}^{-1}$ ).

Figure 4(b) shows the distributions of excitation temperatures  $T_{\text{ex}}$ . The excitation temperatures are generally low, all less than  $50 \text{ K}$ , with a minimum value  $\lesssim 4.5 \text{ K}$ . A vast majority of sources ( $\sim 94\%$ ) show  $T_{\text{ex}}$  less than  $20 \text{ K}$ . The median value is  $\sim 11.7 \text{ K}$ , which is slightly less than that of all infall sources in the literature (Yu et al. 2022). The distribution of Pair-2 does not seem to follow a regular function, probably because of the paucity of samples. On the other hand, the profiles of Pair-1 as well as of the overall are quite smooth. The best log-normal fitting to the overall distribution gives a parameter-pair  $(\sigma, \mu) = (2.41, 0.25)$  with  $\chi^2 = 3.2$ . This result is very close to that derived from Planck cold clumps (Wu et al. 2012), indicating that our sources may be of similar properties to those cold clumps. We note however, as stated in Section 3, the excitation temperatures derived from this work are subject to rather large uncertainties.

Figure 4(c) presents the distributions of column densities of molecular hydrogen,  $N(\text{H}_2)$ . It can be clearly seen the column

densities traced by  $^{13}\text{CO}$  Pair-1 and  $\text{C}^{18}\text{O}$  (Pair-2) are obviously separated. The values for the former group range from  $2.6 \times 10^{20}$  to  $1.1 \times 10^{23} \text{ cm}^{-2}$  with a median value of  $1.3 \times 10^{21} \text{ cm}^{-2}$ , while those for the latter range between  $1.1 \times 10^{21}$  and  $1.8 \times 10^{23} \text{ cm}^{-2}$  with a median value of  $6.4 \times 10^{21} \text{ cm}^{-2}$ . The column densities of our samples are distributed narrower, and have smaller median value than those infall sources in the literature (Yu et al. 2022). This may be because the latter utilized a large variety of molecular lines, which include high density tracers, and used different observing facilities including ones with high angular resolution.

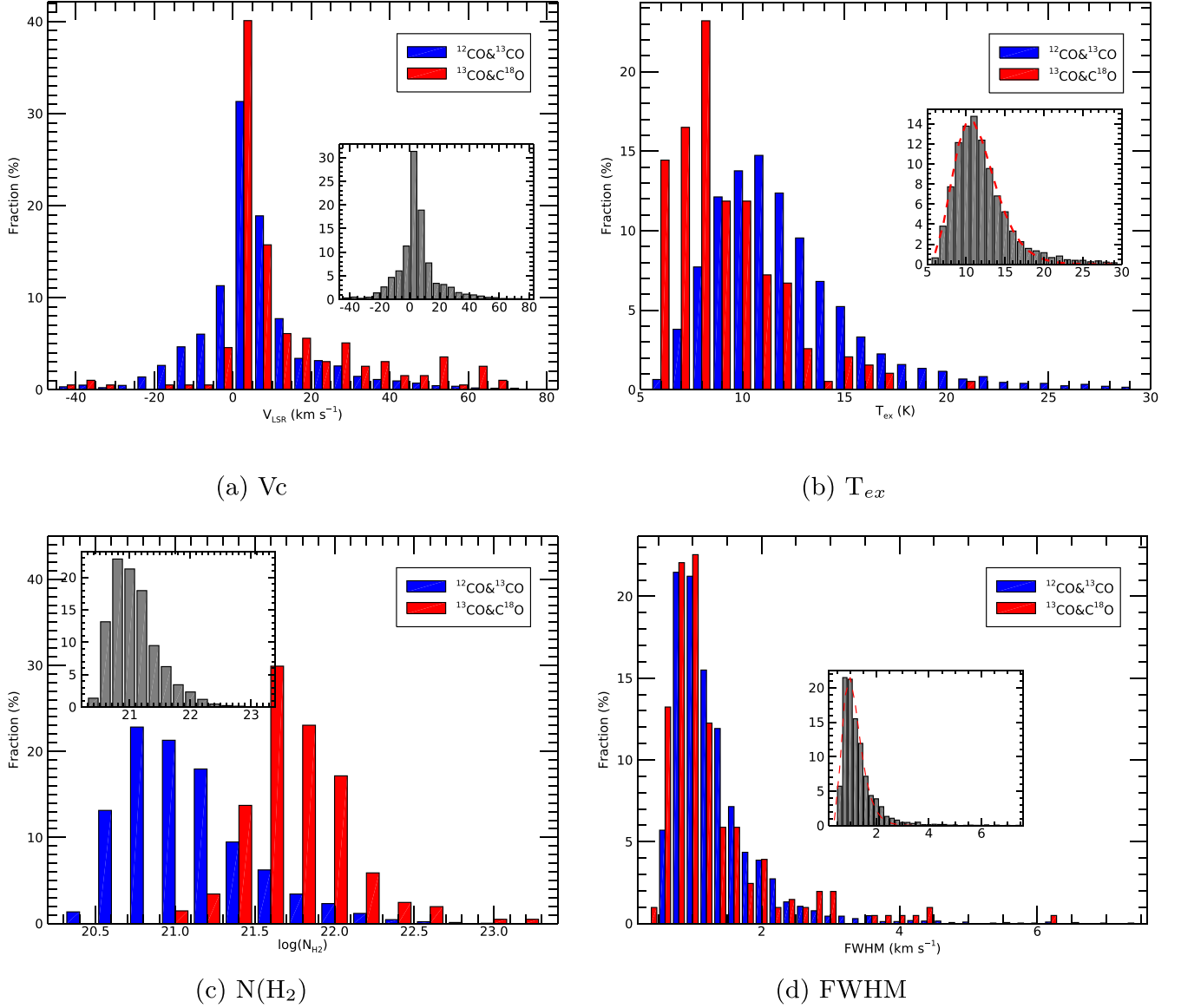
Our result means that a major part of the Pair-2 and a significant part of Pair-1 sources have higher column densities than the possible star formation threshold, i.e.,  $6.3 \times 10^{21} \text{ cm}^{-2}$  (Lada et al. 2010; Kainulainen et al. 2014; Ma et al. 2021). All of our candidates have column densities less than the threshold of massive star formation proposed by Krumholz & McKee (2008), i.e.,  $N(\text{H}_2) \sim 2.1 \times 10^{23} \text{ cm}^{-2}$ . Three factors may reduce the estimated value of  $N(\text{H}_2)$ : (1) large beam size of the telescope (hence low filling factor) may dilute the peak values; (2) our candidates are frequently seen outside of the central part of certain clumps; and (3)  $T_{\text{ex}}$ 's of Pair-2 selected candidates are often underestimated hence the derived  $\text{H}_2$  column densities. Therefore, whether our sources are massive star-forming candidates should await further deliberate studies.

The distribution of the column densities from  $\text{C}^{18}\text{O}$  is roughly symmetric with respect to the median value, suggesting a log-normal profile (note the abscissa is in log-scale). On the other hand, that from  $^{13}\text{CO}$  shows a log-normal style only in the first several bins, up to  $\sim 4.0 \times 10^{21} \text{ cm}^{-2}$ , and deviate from the log-normal distribution at higher column densities. Since Pair-1 objects account for a vast majority of our sample, the overall distribution generally resembles that from  $^{13}\text{CO}$  (inset of Figure 4(c)).

Figure 4(d) shows the distributions of the FWHMs of optically thin lines. The distributions of the sources selected from two pairs do not show significant difference, except for a slightly shift of median values ( $1.21$  versus  $1.10 \text{ km s}^{-1}$  for Pair-1 and Pair-2, respectively). A Kolmogorov–Smirnov (K-S) test with a  $+0.10 \text{ km s}^{-1}$  shift of Pair-2 sources results in  $P$ -value of  $\sim 0.47$ , rejecting the hypothesis that the two profiles are different. A log-normal fit to the overall distribution (inset) results in a  $\sigma = 0.35 \pm 0.013$  and  $\mu = 0.076 \pm 0.014$ . A K-S test results in  $P$ -value of  $0.23$ , thus accepting the log-normal distribution hypothesis.

In Figure 5 we show the distribution of the skewness parameters,  $\delta V$ , defined by Mardones et al. (1997). As can be seen in the figure, the values range between  $-1.6$  and  $0.0$  with a mean value  $\sim -0.5$ . Since our samples are selected in favor of blue-profiles, it is not surprising that the values are mostly less than  $-0.25$ , a criterion as significant infall signature suggested by Mardones et al. (1997).





**Figure 4.** Frequency distributions of (a) central velocities; (b) excitation temperatures; (c) derived  $\text{H}_2$  column densities ( $\text{cm}^{-2}$  on log scale); (d) line widths (FWHM) of the optically thin lines. In the inset of (b), the dashed curve is a log-normal fit with  $(\sigma, \mu) = (2.41, 0.25)$  to overall  $T_{\text{ex}}$  distribution. In the inset of (d), the dashed curve is the best log-normal fit with  $(\sigma, \mu) = (0.34, 0.076)$ , to the overall distribution of FWHMs. For each parameter, the bins for both Pair-1 and Pair-2 as well as for the overall are set to same.

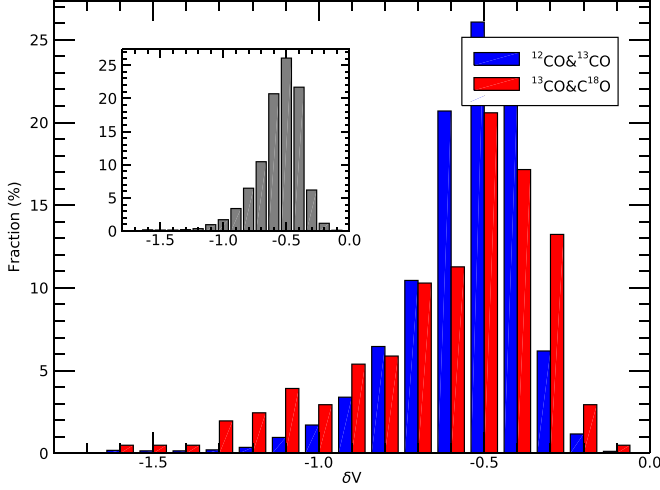
Though  $\delta V$  is defined as inversely correlated to the line widths of the optically thin line, as shown in Figure 6, we have not found any clear trend of such correlation.

#### 4.3. Spatial Distribution

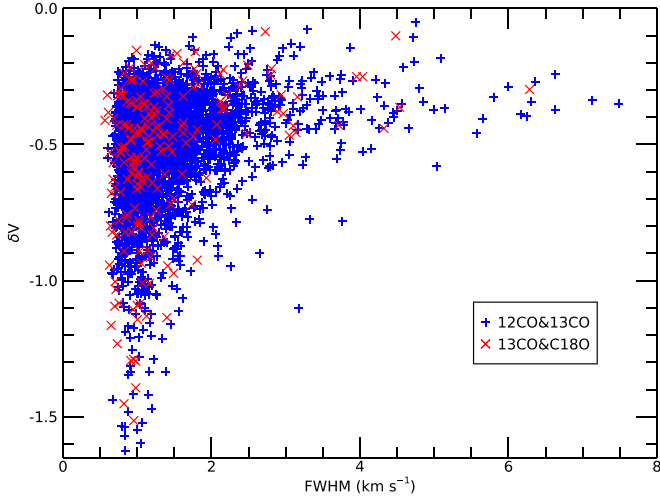
Figure 7 shows the spatial distribution of our candidates, overlaid on the imaginary face-on view of the Milky Way (credit: Xing-Wu Zheng & Mark Reid BeSSeL/NJU/CFA<sup>6</sup>). It

can be seen clearly that a vast majority of our candidates are located in the first quadrant, especially those selected from Pair-2, and outside of the 3 kpc molecular ring. This is not surprising because of the facility limitation and the confusion in the inner Galactic molecular clouds. It should be noted here that most of the sources are located in the spiral arms. This may be due to the distance estimator (Reid et al. 2016, 2019), which assumes that molecular clouds are presumably located in the spiral arms. Though this could be true for a majority of the our sources, we advice the researchers to interpret this result with caution.

<sup>6</sup> <https://astronomy.nyu.edu.cn/xtzl/EN/index.html>



**Figure 5.** The distributions of the skewness parameters,  $\delta V$ , of the sources selected from Pair-1 (blue), Pair-2 (red) and overall (inset).



**Figure 6.** A plot of  $\delta V$  and line width (FWHM) of all candidates. The blue pluses and red crosses represent Pair-1 and Pair-2 selected sources, respectively.

Figure 8 is the vertical distribution (source offset from the Galactic midplane) of our candidates. The main frame shows the pair-separated distribution while the inset is that of overall. Gaussian fitting of the overall distribution shows the FWHM thickness is  $\sim 85$  pc, which is reminiscent of the thin molecular disk suggested by Su et al. (2021). A majority (3034%,  $\sim 85.8\%$ ) of the sources are located within this range (i.e.,  $|z| \leq 42.5$  pc). The profile of Pair-1 selected candidates is wider than that of the Pair-2 ones (85 versus 54 pc in FWHM). Interestingly, we notice the presence of excess of sources at the foot of the Gaussian fit (inset of Figure 8), similar to that of the vertical distribution of molecular clouds by Su et al. (2021), and suggesting that star formation may also take place in the

thick disk component. It is also interesting to mention that a number (16) of sources are located far beyond the Galactic midplane (i.e.,  $|z| \geq 200$  pc). These sources are worthy to investigate to show if there are star-forming activities even far from the Galactic plane.

In Figure 9 we show source distribution in  $L$ - $V_{\text{LSR}}$  space. Apart from the fact that sources selected from Pair-2 are more concentrated to the first quadrant, the overall distribution is reminiscent of the result by Dame et al. (2001, Figure 3). Our sources exist in virtually all large-scale components in the northern galaxy, such as the Sagittarius-Scutum arm, the 3 kpc molecular ring, Lindblad Ring and the Local Arm, the Perseus arm, and even the outer arm. If our sources represent quite a part of the infalling gas motion, the overall distribution indicates that there exist star-forming activities in nearly all large components in the galaxy at present epoch.

#### 4.4. Confirmation Rates

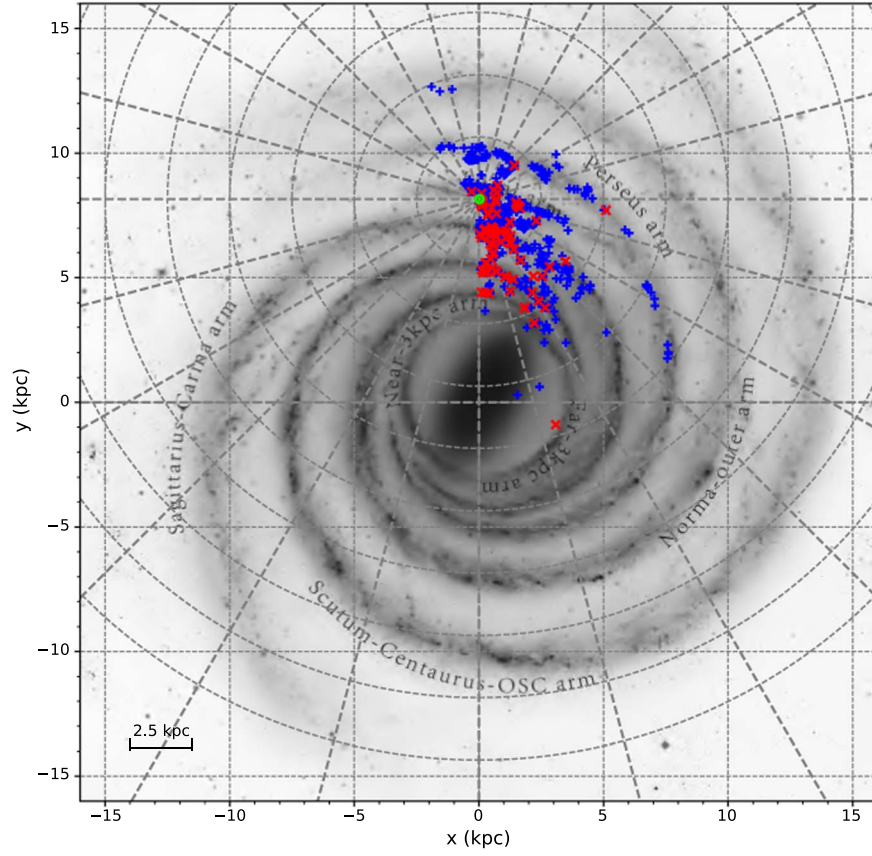
As known to researchers,  $^{12}\text{CO}$  and its isotopic molecules are not ideal tracers to the infall motions, and  $\text{HCO}^+$  would be much more efficient. Therefore, shortly after we initiated this project, follow-up observations were made with  $\text{HCO}^+$  ( $J = 1-0$ ) and  $\text{HCN}$  ( $J = 1-0$ ) lines. Yang et al. (2020) selected 133 candidates from our sample with a restriction that  $T_{\text{MB}}(\text{C}^{18}\text{O}) \geq 1$  K. They confirmed 56 sources to have infall signatures in  $\text{HCO}^+$  and/or  $\text{HCN}$  lines. The overall confirmation rate is  $\sim 42\%$ , the respective rate being 40% and 49% for Pair-1 and Pair-2. Further, Yang et al. (2021) used IRAM to carry out mapping observations toward 24 out of the 56 confirmed sources and finally confirmed nine blue asymmetry profiles in  $\text{HCO}^+$  ( $J = 1-0$ ) line. These nine sources are regarded as infall sources with high confidence.

On the other hand, S. Yu et al. (2023, in preparation) selected a sample of 210 sources with more relaxed conditions to study the infall signatures. They found a total of 40 sources showing blue-profiles in  $\text{HCO}^+$  emission line. The gross detection rate is  $\sim 19\%$ , with 18% for Pair-1 and 28% for Pair-2 objects. In general, Pair-2 selected candidates have higher confirmation rate than Pair-1 selected ones. This is not surprising in light of that  $\text{C}^{18}\text{O}$  traces higher column density than  $^{13}\text{CO}$  does. If the sub-sample represents our overall sample adequately, we would expect the confirmation rate being no less  $\sim 20\%$ .

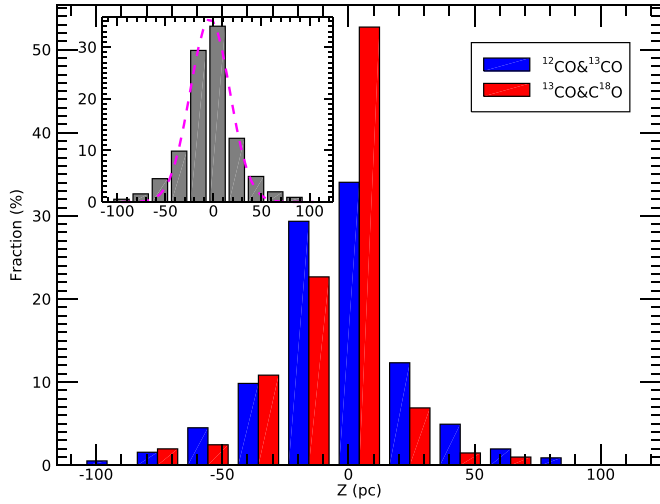
In conclusion, starting from the  $^{12}\text{CO}$ ,  $^{13}\text{CO}$  and  $\text{C}^{18}\text{O}$  ( $J = 1-0$ ) lines, our approach to search for infall sources is effective.

#### 5. Summary

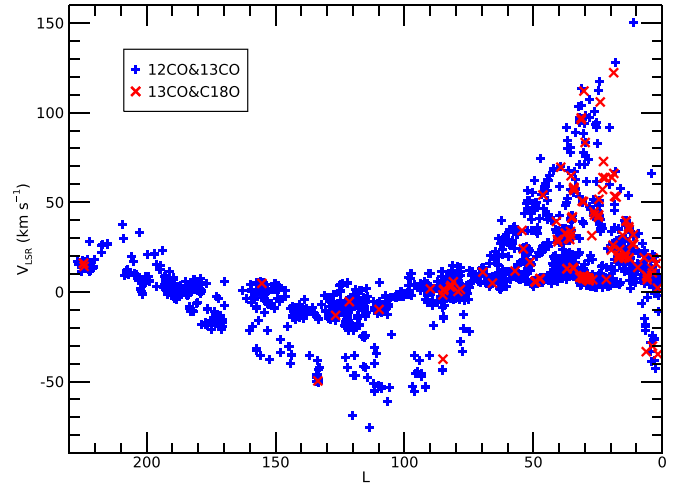
Based on the MWISP data, we conducted a survey of infalling clumps in the galaxy, which presumably show blue asymmetric profiles in the optically thick lines and single peaks in the optically thin lines. Both  $^{12}\text{CO}$ & $^{13}\text{CO}$  and  $^{13}\text{CO}$ & $\text{C}^{18}\text{O}$  pairs are utilized to carry out the work. The automatic search



**Figure 7.** Spatial distribution of the infall candidates overlaid on the imaginary face-on view of the Milky Way (credit: Xing-Wu Zheng & Mark Reid BeSSeL/NJU/CFA). The green  $\odot$  sign indicates the location of the Sun. The Galactic center is at the origin. Our candidates are designated by blue pluses (Pair-1) and red crosses (Pair-2).



**Figure 8.** The vertical distribution. The blue and red bars represent the Pair-1 and Pair-2 selected sources, respectively. The pink dashed curve in the inset is a Gaussian fit to the overall distribution with  $\sigma = 36$  pc (FWHM = 85 pc).



**Figure 9.** The source distribution in  $L$ - $V_{\text{LSR}}$  space. As in Figure 7, our sources are designated as the blue pluses (Pair-1) and red crosses (Pair-2).

and manual check are conducted to select candidates out of  $\sim 10^8$  spectra. The main results are summarized as follows:

1. A total of 3533 sources are finalized as candidates in the infalling phase of star formation, in which 3329 candidates are selected from the  $^{12}\text{CO}$ & $^{13}\text{CO}$  pair and 204 are from the  $^{13}\text{CO}$ & $\text{C}^{18}\text{O}$  pair.
2. Though being manually checked and filtered, the candidates show a wide range of spectra with complicated profiles. The locations that show blue-profiles do not always coincide with central parts of molecular clumps where the highest column densities are detected, but a significant part are rather located at the edges.
3. The analysis of physical parameters of the sources suggests the Pair-2 candidates are colder, and have higher column densities, than Pair-1 ones. The overall distribution of  $T_{\text{ex}}$  follows a log-normal style and is consistent to that by Wu et al. (2012), suggesting the properties of our sources are quite similar to those of Planck cold clumps.
4. The line widths of the optically thin lines for both Pair-1 and Pair-2 sources are in log-normal style. The mean values of the two groups are different by  $< 0.1 \text{ km s}^{-1}$ . A K-S test indicates that the distributions are the same when line widths of Pair-2 sources are added by  $0.1 \text{ km s}^{-1}$ .
5. Most of the sources are located within the first quadrant, especially those selected from the  $^{13}\text{CO}$ & $\text{C}^{18}\text{O}$  pair. The system velocities of the candidates range from  $\sim -70$  to  $\sim 150 \text{ km s}^{-1}$ , and are present in virtually all large scale components in the northern galaxy. The vertical

distribution suggests that our sources are located primarily within the thin disk, but still present in the extended thick disk.

6. A sketchy estimation suggests that the confirmation rate of our sample could be no less than  $\sim 20\%$ , indicating our strategy is a good start to study the very early phase of star formation systematically.

### Acknowledgments

We are deeply obliged to the MWISP working group, Yang Su, Xin Zhou, Yan Sun, Jixian Sun, Dengrong Lu and Binggang Ju, and the observation assistants of the project, who have been working hard in instrument maintenance, taking and reducing the data, without which this work cannot be done. This work is supported by the National Key R&D Program of China (Grant No. 2017YFA0402702), and the National Natural Science Foundation of China (NSFC, Grant Nos., 11873093, U2031202, and 11903083). MWISP is sponsored by the National Key R&D Program of China with Grant 2017YFA0402701 and by CAS Key Research Program of Frontier Sciences with grant QYZDJ-SSW-SLH047. We would like to express our thanks to the anonymous reviewer for constructive suggestions that enable improvement of the manuscript.

### Appendix A The Full Catalogue

In Table A1 we present the full catalog of all candidates. The descriptions of the columns are described in the main text.



**Table A1**  
The Catalog

No	Code	$L$	$B$	Association <sup>a</sup>	I.S. <sup>a</sup>	$D^b$ (kpc)	$V_c$ (km s <sup>-1</sup> )	$V_b$ (km s <sup>-1</sup> )	$\delta V$	$T_1$ (K)	$T_2$ (K)	$T_{\text{ex}}$ (K)	$N_{\text{H}_2}^c$ (cm <sup>-2</sup> )	Pair
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
1001	03 105+0017+0118	31.0500	0.1667	1, 3, 6	...	1.90	11.83	11.20	-0.65	4.1	1.2	11.2	5.04E+20	1
1002	03 105+0076+0506	31.0500	0.7583	6, 7	...	0.68	50.62	50.07	-0.44	4.2	1.3	8.8	5.82E+21	2
1003	03 105+0407+0063	31.0500	4.0750	5, 6	...	0.24	6.29	5.68	-0.59	5.8	1.2	8.4	5.27E+20	1
1004	03 106-0067+0088	31.0583	-0.6667	6, 7	...	0.24	8.77	8.41	-0.42	4.1	1.8	8.3	7.75E+20	1
1005	03 107+0078+0512	31.0667	0.7833	6, 7	...	0.60	51.22	50.26	-0.48	6.5	2.9	11.2	3.02E+21	1
1006	03 107+0313+0103	31.0750	3.1333	5, 6	...	0.24	10.31	9.32	-0.46	6.8	3.4	10.0	4.08E+21	1
1007	03 107+0524+0084	31.0750	5.2417	5, 6	...	0.24	8.41	7.40	-0.51	2.7	3.4	6.9	6.93E+21	1
1008	03 107-0044+0085	31.0750	-0.4417	6	...	0.24	8.50	8.14	-0.32	4.3	1.3	10.9	6.56E+20	1
1009	03 108-0042+0086	31.0833	-0.4167	6	...	0.24	8.61	7.96	-0.68	4.3	1.3	9.1	5.32E+20	1
1010	03 109-0339+0100	31.0917	-3.3917	5, 6	...	0.24	10.00	9.83	-0.23	5.2	2.1	10.3	7.41E+20	1

**Notes.** Column Heads: (1): internal serial number; (2): name code; (3) and (4): Galactic Coordinates; (5): association to some known catalogs (see, next foot note); (6): infall signature reported in the literature (see next foot note); (7): distance; (8): central velocity; (9): velocity at the blue peak; (10): skewness parameter (see text); (11): expected Tmb of optically thick line; (12): peak Tmb of optically thin line; (13): excitation temperature; (14): H<sub>2</sub> column density; (15): pair from which the blue-profile is obtained.

<sup>a</sup> Ref. Codes: (1) IRAS Point Source Catalog; (2) New General Catalog (NGC); (3) Price & Murdock (1999); (4) Sharpless (1953); (5) Lynds (1996); (6) Anderson et al. (2014); (7) Simon et al. (2006); (8) Yan et al. (2018); (9) Shirley et al. (2003); (10) Fuller et al. (2005); (11) Klaassen & Wilson (2007); (12) Sun & Gao (2009); (13) Reiter et al. (2011); (14) Klaassen et al. (2012); (15) Liu et al. (2013); (16) Qin et al. (2016); (17) He et al. (2016); (18) Yang et al. (2020); (19) Kim et al. (2021).

<sup>b</sup> For a few sources the distances are not given because the distance calculator does not output proper values.

<sup>c</sup> For a number of entries the values of N(H<sub>2</sub>) are not given because the optical depths ( $\tau_{\text{thin}}$ ) cannot be estimated due to the complex line profiles.

## Appendix B

Spectra and gas distributions of all candidates. symbols are the same as in Figure 2 in the main text. Source name is indicated on the upper-left corner of each panel.

## ORCID iDs

Zhibo Jiang (江治波)  <https://orcid.org/0000-0002-5920-031X>  
 Shaobo Zhang (张少博)  <https://orcid.org/0000-0003-2549-7247>  
 Zhiwei Chen (陈志维)  <https://orcid.org/0000-0003-0849-0692>  
 Yang Yang (杨旸)  <https://orcid.org/0000-0002-2895-7707>  
 Shuling Yu (于书岭)  <https://orcid.org/0000-0002-8617-3521>  
 Haoran Feng (冯浩然)  <https://orcid.org/0000-0003-1714-0600>  
 Ji Yang (杨戟)  <https://orcid.org/0000-0001-7768-7320>

## References

- Anderson, L. D., Bania, T. M., Balser, D. S., et al. 2014, *ApJS*, **212**, 1  
 Areal, M. B., Paron, S., Ortega, M. E., & Duvidovich, L. 2019, *PASA*, **36**, e049  
 Calahan, J. K., Shirley, Y. L., Svoboda, B. E., et al. 2018, *ApJ*, **862**, 63  
 Chen, X., Shen, Z.-Q., Li, J.-J., Xu, Y., & He, J.-H. 2010, *ApJ*, **710**, 150  
 Dame, T. M., Hartmann, D., & Thaddeus, P. 2001, *ApJ*, **547**, 792  
 Evans, N. J. I. 1999, *ARA&A*, **37**, 311  
 Fuller, G. A., Williams, S. J., & Sridharan, T. K. 2005, *A&A*, **442**, 949  
 He, Y.-X., Zhou, J.-J., Esimbek, J., et al. 2015, *MNRAS*, **450**, 1926  
 He, Y.-X., Zhou, J.-J., Esimbek, J., et al. 2016, *MNRAS*, **461**, 2288  
 Heyer, M., & Dame, T. M. 2015, *ARA&A*, **53**, 583  
 Jin, M., Lee, J.-E., Kim, K.-T., & II, N. J. E. 2016, *ApJS*, **225**, 21  
 Kainulainen, J., Federrath, C., & Henning, T. 2014, *Science*, **344**, 183  
 Kim, M.-R., Lee, C. W., Maheswar, G., Myers, P. C., & Kim, G. 2021, *ApJ*, **910**, 112  
 Klaassen, P. D., Testi, L., & Beuther, H. 2012, *A&A*, **538**, A140  
 Klaassen, P. D., & Wilson, C. D. 2007, *ApJ*, **663**, 1092  
 Krumholz, M. R., & McKee, C. F. 2008, *Natur*, **451**, 1082  
 Lada, C. J., Lombardi, M., & Alves, J. F. 2010, *ApJ*, **724**, 687  
 Liu, T., Wu, Y., & Zhang, H. 2013, *ApJ*, **776**, 29  
 Lynds, B. T. 1996, *yCat*, **VII/7A**  
 Ma, Y., Wang, H., Li, C., et al. 2021, *ApJS*, **254**, 3  
 Mangum, J. G., & Shirley, Y. L. 2015, *PASP*, **127**, 266  
 Mangum, J. G., & Shirley, Y. L. 2016, *PASP*, **128**, 029201  
 Mardones, D., Myers, P., Caselli, P., & Fuller, G. 1994, in ASP Conf. Ser. 65, Clouds, Cores, and Low Mass Stars, ed. D. P. Clemens & R. Barvainis (San Francisco, CA: ASP), 192  
 Mardones, D., Myers, P. C., Tafalla, M., et al. 1997, *ApJ*, **489**, 719  
 Martin, R. N., & Barrett, A. H. 1978, *ApJS*, **36**, 1  
 Mayen-Gijon, J. M., Anglada, G., Osorio, M., et al. 2014, *MNRAS*, **437**, 3766  
 Myers, P. C., Mardones, D., Tafalla, M., Williams, J. P., & Wilner, D. J. 1996, *ApJL*, **465**, L133  
 Pineda, J. E., Caselli, P., & Goodman, A. A. 2008, *ApJ*, **679**, 481  
 Price, S. D., & Murdock, T. L. 1999, *yCat*, **II/94**  
 Qin, S.-L., Schilke, P., Wu, J., et al. 2016, *MNRAS*, **456**, 2681  
 Reid, M. J., Dame, T. M., Menten, K. M., & Brunthaler, A. 2016, *ApJ*, **823**, 77  
 Reid, M. J., Menten, K. M., Brunthaler, A., et al. 2019, *ApJ*, **885**, 131  
 Reiter, M., Shirley, Y. L., Wu, J., et al. 2011, *ApJ*, **740**, 40  
 Ren, Z., Wu, Y., Zhu, M., et al. 2012, *MNRAS*, **422**, 1098  
 Rygl, K. L. J., Wyrowski, F., Schuller, F., & Menten, K. M. 2013, *A&A*, **549**, A5  
 Shan, W., Yang, J., Shi, S., et al. 2012, *ITST*, **2**, 593  
 Sharpless, S. 1953, *ApJ*, **118**, 362  
 Shirley, Y. L., Evans, N. J. I., Young, K. E., Knez, C., & Jaffe, D. T. 2003, *ApJS*, **149**, 375  
 Shu, F. H., Adams, F. C., & Lizano, S. 1987, *ARA&A*, **25**, 23  
 Simon, R., Jackson, J. M., Rathborne, J. M., & Chambers, E. T. 2006, *ApJ*, **639**, 227  
 Su, Y., Yang, J., Yan, Q.-Z., et al. 2021, *ApJ*, **910**, 131  
 Su, Y., Yang, J., Zhang, S., et al. 2019, *ApJS*, **240**, 9  
 Sun, Y., & Gao, Y. 2009, *MNRAS*, **392**, 170  
 Wu, J., & Evans, N. J., II 2003, *ApJL*, **592**, L79  
 Wu, Y., Henkel, C., Xue, R., Guan, X., & Miller, M. 2007, *ApJL*, **669**, L37  
 Wu, Y., Liu, T., Meng, F., et al. 2012, *ApJ*, **756**, 76  
 Yan, Q.-Z., Xu, Y., Walsh, A. J., et al. 2018, *MNRAS*, **476**, 3981  
 Yang, Y., Jiang, Z., Chen, Z., Ao, Y., & Yu, S. 2021, *ApJ*, **922**, 144  
 Yang, Y., Jiang, Z.-B., Chen, Z.-W., et al. 2020, *RAA*, **20**, 115  
 Yoo, H., Kim, K.-T., Cho, J., et al. 2018, *ApJS*, **235**, 31  
 Yu, S., Jiang, Z., Yang, Y., Chen, Z., & Feng, H. 2022, *RAA*, **22**, 095014  
 Yuan, J., Li, J.-Z., Wu, Y., et al. 2018, *ApJ*, **852**, 12  
 Zhou, S., Evans, N. J., II, Koempe, C., & Walmsley, C. M. 1993, *ApJ*, **404**, 232  
 Zhu, L., Zhao, J.-H., & Wright, M. C. H. 2011, *ApJ*, **740**, 114